Do-calculus

Charles K. Assaad, Emilie Devijver, Eric Gaussier

charles.assaad@ens-lyon.fr

Table of content

Table of content

Active and blocked paths A path is said to be blocked by a set of vertices $\mathcal{Z} \in \mathcal{V}$ if:

- it contains a chain $A \rightarrow B \rightarrow C$ or a fork $A \leftarrow B \rightarrow C$ and $B \in \mathcal{Z}$, or
- it contains a collider $A \rightarrow B \leftarrow C$ such that no descendant of B is in $\mathcal{Z}$.

d-separation Given disjoint sets $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \subseteq \mathcal{V}$, we say that $\mathcal{X}$ and $\mathcal{Y}$ are d-separated by $\mathcal{Z}$ if every path between a node in $\mathcal{X}$ and a node in $\mathcal{Y}$ is blocked by $\mathcal{Z}$ and we write $\mathcal{X} \perp\!\!\!\perp_G \mathcal{Y} | \mathcal{Z}$.

Recap about causal graphical models (2/2)

The do() operator allows to represent interventions in equations.

Recap about the Back-door and Front-door criteria (1/3)

The back-door criterion: Consider a causal graph $\mathcal{G}$ and a causal effect $P(y \mid do(x))$. A set of variables $\mathcal{Z}$ satisfies the back-door criterion iff:

- no node in $\mathcal{Z}$ is a descendant of X;
- $\mathcal{Z}$ blocks every path between X and Y that contains an arrow into X.

Theorem (back-door adjustment): If $\mathcal{Z}$ satisfies the back-door criterion relative to (X, Y) and if $\Pr(x, z) > 0$, then the causal effect of X on Y is identifiable and is given by

$$\Pr(y \mid do(x)) = \sum_z \Pr(y \mid x, z) \Pr(z).$$

Recap about the Back-door and Front-door criteria (2/3)

Front-door criterion: Consider a causal graph $\mathcal{G}$ and a causal effect $\Pr(y \mid do(x))$. A set of variables $\mathcal{Z}$ satisfies the front-door criterion iff:

- $\mathcal{Z}$ intercepts all directed paths from X to Y;
- There is no back-door path from X to $\mathcal{Z}$;
- All back-door paths from $\mathcal{Z}$ to Y are blocked by X.

Theorem (front-door adjustment): if $\mathcal{Z}$ satisfies the front-door criterion relative to (X, Y) and if $\Pr(x, z) > 0$, then the causal effect of X on Y is identifiable and is given by

$$\Pr(y \mid do(X = x)) = \sum_{z} \Pr(z \mid x) \sum_{x'} \Pr(y \mid x', z) \Pr(x').$$

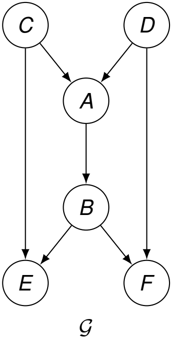
Recap about the Back-door and Front-door criteria (3/3)

- If there exists a set that satisfy the back-door criterion for $\Pr(y \mid do(x))$, then $\Pr(y \mid do(x))$ is identifiable;
- If there exists a no set that satisfy the back-door criterion for $\Pr(y \mid do(x))$, then $\Pr(y \mid do(x))$ is not necesarly not identifiable.
- If there exists a set that satisfy the front-door criterion for $\Pr(y \mid do(x))$, then $\Pr(y \mid do(x))$ is identifiable;
- If there exists a no set that satisfy the fack-door criterion for $\Pr(y \mid do(x))$, then $\Pr(y \mid do(x))$ is not necesarly not identifiable.

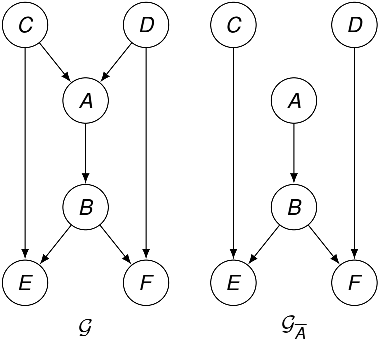
The combination of the back-door and front door criteria are also incomplete.

Table of content

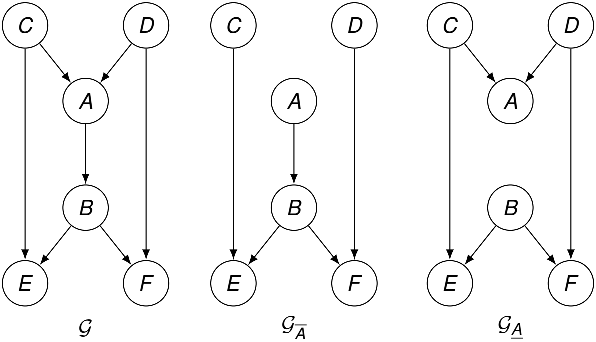
Mutilated Graphs



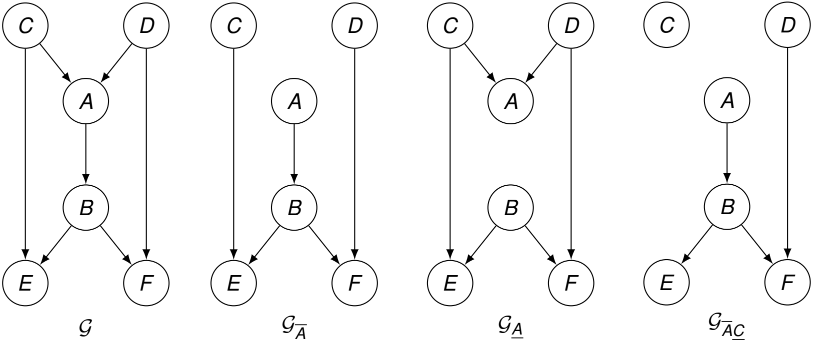
Mutilated Graphs



Mutilated Graphs



Mutilated Graphs



Augmented Graphs

Consider $\Pr(y \mid do(z))$ and the Probabilistic Causal Model:

$$M = < \mathcal{U}, \mathcal{V}, \mathcal{F}, P(\mathcal{U}) >$$

Augmented Graphs

Consider $\Pr(y \mid do(z))$ and the Probabilistic Causal Model:

$$M = < \mathcal{U}, \mathcal{V}, \mathcal{F}, P(\mathcal{U}) >$$

Augmented model of M for $do(z)$

$$Aug(M, \mathcal{Z}) = < \mathcal{U}, \mathcal{V} \cup \hat{\mathcal{Z}}, \mathcal{F}_{\hat{\mathcal{Z}}}, P(\mathcal{U}) >$$

where $\forall \hat{Z} \in \hat{\mathcal{Z}}$, $\hat{Z}$ represents $do(z)$.

Augmented Graphs

Consider $\Pr(y \mid do(z))$ and the Probabilistic Causal Model:

$$M = <\mathcal{U}, \mathcal{V}, \mathcal{F}, P(\mathcal{U})>$$

Augmented model of M for $do(z)$

$$Aug(M, \mathcal{Z}) = <\mathcal{U}, \mathcal{V} \cup \hat{\mathcal{Z}}, \mathcal{F}_{\hat{\mathcal{Z}}}, P(\mathcal{U})>$$

where $\forall \hat{Z} \in \hat{\mathcal{Z}}$, $\hat{Z}$ represents $do(z)$.

Augmented graph of $\mathcal{G}$ for $do(z)$

$$Aug(\mathcal{G}, \mathcal{Z}) = \mathcal{G} \cup \{\hat{Z} \rightarrow Z \mid \forall \hat{Z} \in \hat{\mathcal{Z}}\}$$

Augmented Graphs

Consider $\Pr(y \mid do(z))$ and the Probabilistic Causal Model:

$$M =< \mathcal{U}, \mathcal{V}, \mathcal{F}, P(\mathcal{U}) >$$

Augmented model of M for $do(z)$

$$Aug(M, \mathcal{Z}) =< \mathcal{U}, \mathcal{V} \cup \hat{\mathcal{Z}}, \mathcal{F}_{\hat{\mathcal{Z}}}, P(\mathcal{U}) >$$

where $\forall \hat{Z} \in \hat{\mathcal{Z}}$, $\hat{Z}$ represents $do(z)$.

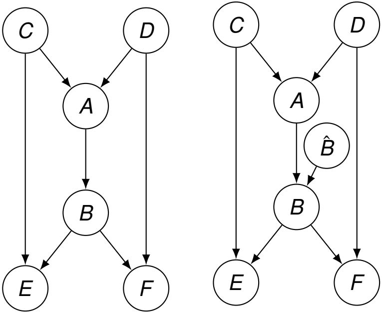
Augmented graph of $\mathcal{G}$ for $do(z)$

$$Aug(\mathcal{G}, \mathcal{Z}) = \mathcal{G} \cup \{\hat{Z} \to Z \mid \forall \hat{Z} \in \hat{\mathcal{Z}}\}$$

and in the compatible distribution, $\forall Z \in \mathcal{Z}$

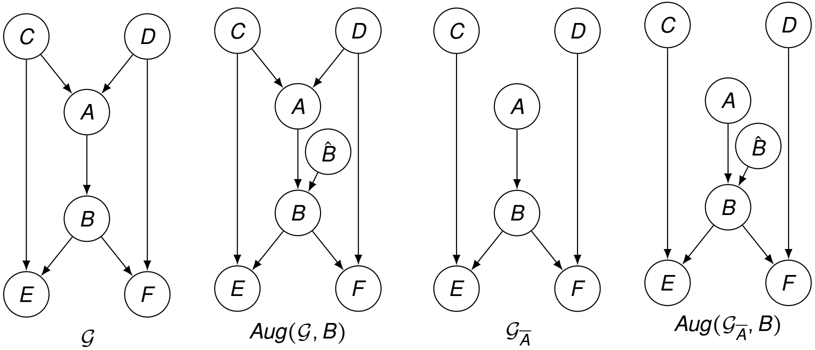
$$P(z \mid Pa(z), \hat{z}) = \begin{cases} P(z \mid Pa(z)) & \text{if } \hat{Z} = idle \\ \hat{z} & \text{if } \hat{Z} = do(z) \end{cases}$$

Example of an augmented graph



$\mathcal{G}$ $\quad$ $Aug(\mathcal{G}, B)$

Example of an augmented graph



Rule 1: Insertion / deletion of observations

Theorem Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a causal graph. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \subseteq \mathcal{V}$ be disjoint. We have:

$$\Pr(y|do(x), z, w) = \Pr(y|do(x), w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\mathcal{X}, \mathcal{W})_{\mathcal{G}_{\overline{\mathcal{X}}}}$$

(proof on board)

Rule 1: Insertion / deletion of observations

Theorem Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a causal graph. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \subseteq \mathcal{V}$ be disjoint. We have:

$$\Pr(y|do(x), z, w) = \Pr(y|do(x), w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\mathcal{X}, \mathcal{W})_{\mathcal{G}_{\overline{\mathcal{X}}}}$$

(proof on board)

Remark: This Rule is a generalization of d-separation.

Rule 2: Action/observation exchange

Theorem Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a causal graph. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \subseteq \mathcal{V}$ be disjoint. We have:

$$\Pr(y | do(x), do(z), w) = \Pr(y | do(x), z, w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z} | \mathcal{X}, \mathcal{W})_{\mathcal{G}_{\overline{\mathcal{X}}\underline{\mathcal{Z}}}}$$

Rule 2: Action/observation exchange

Theorem Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a causal graph. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \subseteq \mathcal{V}$ be disjoint. We have:

$$\Pr(y|do(x), do(z), w) = \Pr(y|do(x), z, w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\mathcal{X}, \mathcal{W})_{\mathcal{G}_{\overline{\mathcal{X}}\underline{\mathcal{Z}}}}$$

Proof: Follows the following Lemma

Lemma Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a causal graph. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \subseteq \mathcal{V}$ be disjoint.

$$(\mathcal{Y} \perp\!\!\!\perp \mathcal{Z} \mid \mathcal{X}, \mathcal{W})_{\mathcal{G}_{\overline{\mathcal{X}}\underline{\mathcal{Z}}}} \iff (\hat{\mathcal{Z}} \perp\!\!\!\perp \mathcal{Y} \mid \mathcal{X}, \mathcal{Z}, \mathcal{W})_{Aug(\mathcal{G}_{\overline{\mathcal{X}}\underline{\mathcal{Z}}}, \mathcal{Z})}$$

(proof on board)

Rule 2: Action/observation exchange

Theorem Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a causal graph. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \subseteq \mathcal{V}$ be disjoint. We have:

$$\Pr(y | do(x), do(z), w) = \Pr(y | do(x), z, w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z} | \mathcal{X}, \mathcal{W})_{\mathcal{G}_{\overline{\mathcal{X}}\underline{\mathcal{Z}}}}$$

Proof: Follows the following Lemma

Lemma Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a causal graph. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \subseteq \mathcal{V}$ be disjoint.

$$(\mathcal{Y} \perp\!\!\!\perp \mathcal{Z} \mid \mathcal{X}, \mathcal{W})_{\mathcal{G}_{\overline{\mathcal{X}}\underline{\mathcal{Z}}}} \iff (\hat{\mathcal{Z}} \perp\!\!\!\perp \mathcal{Y} \mid \mathcal{X}, \mathcal{Z}, \mathcal{W})_{Aug(\mathcal{G}_{\overline{\mathcal{X}}\underline{\mathcal{Z}}}, \mathcal{Z})}$$

(proof on board)

Remark: This Rule is a generalization of the back-door criterion.

Rule 3: insertion / deletion of actions

Theorem Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a causal graph. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \subseteq V$ be disjoint. We have:

$$\Pr(y|do(x), do(z), w) = \Pr(y|do(x), w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\mathcal{X}, \mathcal{W})_{\mathcal{G}_{\overline{\mathcal{X}}\overline{\mathcal{Z}(\mathcal{W})}}}$$

where $\mathcal{Z}(\mathcal{W})$ is the set of $\mathcal{Z}$-vertices that are not ancestors of any $\mathcal{W}$-vertex in $\mathcal{G}_{\overline{\mathcal{X}}}$

Proof in (Pearl, 1995)

Intuition for Rule 3

$$\Pr(y|\cancel{do(x)}, do(z), w) = \Pr(y|\cancel{do(x)}, w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\cancel{\mathcal{X}}, \mathcal{W})_{\mathcal{G}_{\overline{\cancel{\mathcal{X}}}\,\overline{\mathcal{Z}(\mathcal{W})}}}$$

where $\mathcal{Z}(\mathcal{W})$ is the set of $\mathcal{Z}$-vertices that are not ancestors of any $\mathcal{W}$-vertex in $G_{\overline{\cancel{\mathcal{X}}}}$

Intuition for Rule 3

$$\Pr(y|do(z), w) = \Pr(y|w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\mathcal{W})_{\mathcal{G}_{\overline{\mathcal{Z}(\mathcal{W})}}}$$

where $\mathcal{Z}(\mathcal{W})$ is the set of $\mathcal{Z}$-vertices that are not ancestors of any $\mathcal{W}$-vertex in G

Intuition for Rule 3

$$\Pr(y|do(z), w) = \Pr(y|w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\mathcal{W})_{\mathcal{G}_{\overline{\mathcal{Z}(\mathcal{W})}}}$$

where $\mathcal{Z}(\mathcal{W})$ is the set of $\mathcal{Z}$-vertices that are not ancestors of any $\mathcal{W}$-vertex in G

Suppose

$$\Pr(y \mid do(z), w_1, w_2)$$
$$= \Pr(y \mid w_1, w_2)$$
$$\text{if } (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|W_1, W_2)_{\mathcal{G}_{\overline{\mathcal{Z}}}}$$

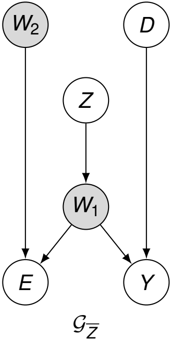
Intuition for Rule 3

$$\Pr(y|do(z), w) = \Pr(y|w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\mathcal{W})_{\mathcal{G}_{\overline{\mathcal{Z}(\mathcal{W})}}}$$

where $\mathcal{Z}(\mathcal{W})$ is the set of $\mathcal{Z}$-vertices that are not ancestors of any $\mathcal{W}$-vertex in G

Suppose

$\Pr(y \mid do(z), w_1, w_2)$
$= \Pr(y \mid w_1, w_2)$
if $(\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|W_1, W_2)_{\mathcal{G}_{\overline{\mathcal{Z}}}}$



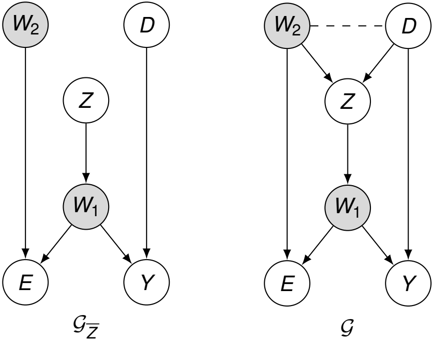
Intuition for Rule 3

$$\Pr(y|do(z), w) = \Pr(y|w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\mathcal{W})_{\mathcal{G}_{\overline{\mathcal{Z}(\mathcal{W})}}}$$

where $\mathcal{Z}(\mathcal{W})$ is the set of $\mathcal{Z}$-vertices that are not ancestors of any $\mathcal{W}$-vertex in G

Suppose

$\Pr(y \mid do(z), w_1, w_2)$
$= \Pr(y \mid w_1, w_2)$
if $(\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|W_1, W_2)_{\mathcal{G}_{\overline{\mathcal{Z}}}}$



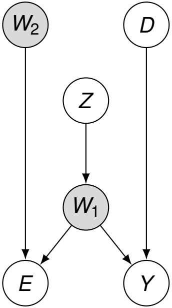
Intuition for Rule 3

$$\Pr(y|do(z), w) = \Pr(y|w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\mathcal{W})_{\mathcal{G}_{\overline{\mathcal{Z}(\mathcal{W})}}}$$

where $\mathcal{Z}(\mathcal{W})$ is the set of $\mathcal{Z}$-vertices that are not ancestors of any $\mathcal{W}$-vertex in G

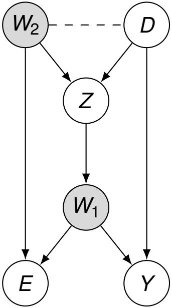
Suppose

$\Pr(y \mid do(z), w_1, w_2)$
$= \Pr(y \mid w_1, w_2)$
if $(\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|W_1, W_2)_{\mathcal{G}_{\overline{\mathcal{Z}}}}$



$\mathcal{G}_{\overline{Z}}$

$Y \perp\!\!\!\perp_{\mathcal{G}_{\overline{Z}}} W_2 \mid W_1$



$\mathcal{G}$

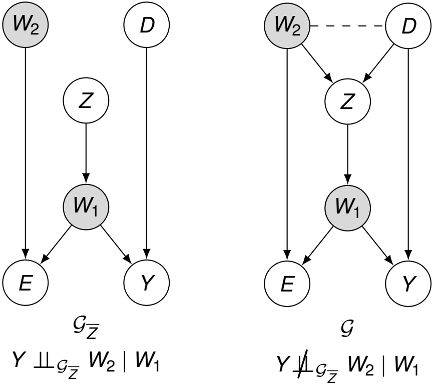
Intuition for Rule 3

$$\Pr(y|do(z), w) = \Pr(y|w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\mathcal{W})_{\mathcal{G}_{\overline{\mathcal{Z}(\mathcal{W})}}}$$

where $\mathcal{Z}(\mathcal{W})$ is the set of $\mathcal{Z}$-vertices that are not ancestors of any $\mathcal{W}$-vertex in G

Suppose

$\Pr(y \mid do(z), w_1, w_2)$
$= \Pr(y \mid w_1, w_2)$
if $(\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|W_1, W_2)_{\mathcal{G}_{\overline{\mathcal{Z}}}}$



$$\Pr(y|do(z), w) = \Pr(y|w) \quad \text{if} \quad (\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|\mathcal{W})_{\mathcal{G}_{\overline{\mathcal{Z}(\mathcal{W})}}}$$

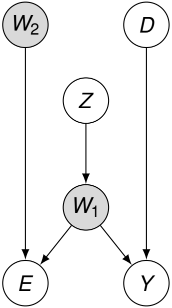
where $\mathcal{Z}(\mathcal{W})$ is the set of $\mathcal{Z}$-vertices that are not ancestors of any $\mathcal{W}$-vertex in G

~~Suppose~~

~~$\Pr(y \mid do(z), w_1, w_2)$~~

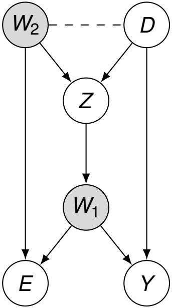
~~$= \Pr(y \mid w_1, w_2)$~~

~~if $(\mathcal{Y} \perp\!\!\!\perp \mathcal{Z}|W_1, W_2)_{\mathcal{G}_{\overline{\mathcal{Z}}}}$~~



$\mathcal{G}_{\overline{Z}}$

$Y \perp\!\!\!\perp_{\mathcal{G}_{\overline{Z}}} W_2 \mid W_1$



$\mathcal{G}$

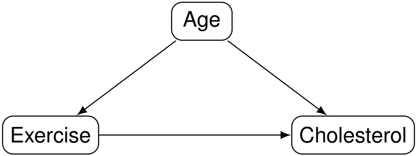
$Y \not\perp\!\!\!\perp_{\mathcal{G}_{\overline{Z}}} W_2 \mid W_1$

Completness of the do-calculus

Theorem A causal effect $P(y \mid do(x))$ is identifiable in a model characterized by a graph $\mathcal{G}$ if and only if there exists a finite sequence of transformations, each conforming to one the Rules 1-3, that reduces $P(y \mid do(x))$ into a standard (i.e., "do"-free) probability expression involving observed quantities.

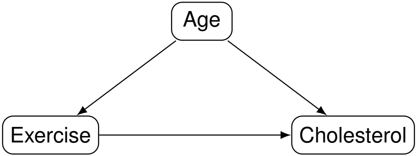
Proof in (Pearl, 1995) and (Shpitser and Pearl, 2006)

From do-calculus to back-door adjustment



What's the effect of exercice on cholesterol?

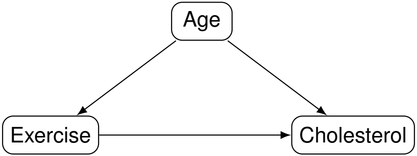
From do-calculus to back-door adjustment



What's the effect of exercice on cholesterol?

$$\Pr(c \mid do(e)) = \sum_a \Pr(c \mid do(e), a) \Pr(a \mid do(e)) \qquad \text{(Probability Axioms)}$$

From do-calculus to back-door adjustment

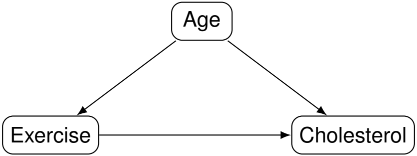


What's the effect of exercice on cholesterol?

$$\Pr(c \mid do(e)) = \sum_a \Pr(c \mid do(e), a) \Pr(a \mid do(e)) \qquad \text{(Probability Axioms)}$$

$$= \sum_a \Pr(c \mid e, a) \Pr(a \mid do(e)) \qquad \text{(Rule 2)}$$

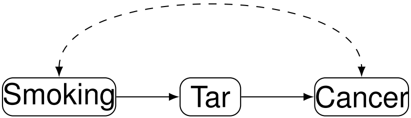
From do-calculus to back-door adjustment



What's the effect of exercice on cholesterol?

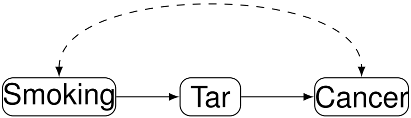
$$\begin{aligned}
\Pr(c \mid do(e)) &= \sum_a \Pr(c \mid do(e), a) \Pr(a \mid do(e)) && \text{(Probability Axioms)} \\
&= \sum_a \Pr(c \mid e, a) \Pr(a \mid do(e)) && \text{(Rule 2)} \\
&= \sum_a \Pr(c \mid e, a) \Pr(a) && \text{(Rule 3)}
\end{aligned}$$

From do-calculus to front-door adjustment



What's the effect of smoking on cancer?

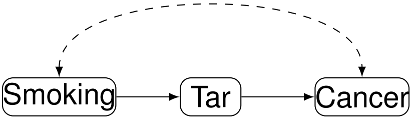
From do-calculus to front-door adjustment



What's the effect of smoking on cancer?

$$\Pr(c \mid do(s)) = \sum_t \Pr(c \mid do(s), t) \Pr(t \mid do(s)) \qquad \text{(Probability Axioms)}$$

From do-calculus to front-door adjustment

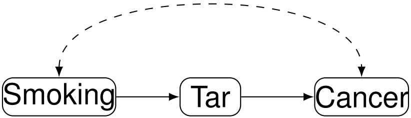


What's the effect of smoking on cancer?

$$\Pr(c \mid do(s)) = \sum_t \Pr(c \mid do(s), t) \Pr(t \mid do(s)) \quad \text{(Probability Axioms)}$$

$$= \sum_t \Pr(c \mid do(s), do(t)) \Pr(t \mid do(s)) \quad \text{(Rule 2)}$$

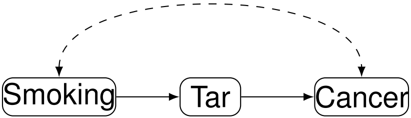
From do-calculus to front-door adjustment



What's the effect of smoking on cancer?

$$
\begin{aligned}
\Pr(c \mid do(s)) &= \sum_t \Pr(c \mid do(s), t) \Pr(t \mid do(s)) && \text{(Probability Axioms)} \\
&= \sum_t \Pr(c \mid do(s), do(t)) \Pr(t \mid do(s)) && \text{(Rule 2)} \\
&= \sum_t \Pr(c \mid do(s), do(t)) \Pr(t \mid s) && \text{(Rule 2)}
\end{aligned}
$$

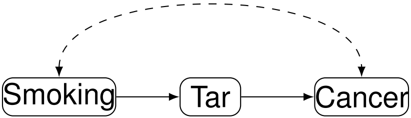
From do-calculus to front-door adjustment



What's the effect of smoking on cancer?

$$
\begin{aligned}
\Pr(c \mid do(s)) &= \sum_t \Pr(c \mid do(s), t) \Pr(t \mid do(s)) && \text{(Probability Axioms)} \\
&= \sum_t \Pr(c \mid do(s), do(t)) \Pr(t \mid do(s)) && \text{(Rule 2)} \\
&= \sum_t \Pr(c \mid do(s), do(t)) \Pr(t \mid s) && \text{(Rule 2)} \\
&= \sum_t \Pr(c \mid do(t)) \Pr(t \mid s) && \text{(Rule 3)}
\end{aligned}
$$

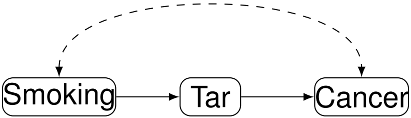
From do-calculus to front-door adjustment



What's the effect of smoking on cancer?

$$
\begin{aligned}
\Pr(c \mid do(s)) &= \sum_t \Pr(c \mid do(s), t)\Pr(t \mid do(s)) && \text{(Probability Axioms)}\\
&= \sum_t \Pr(c \mid do(s), do(t))\Pr(t \mid do(s)) && \text{(Rule 2)}\\
&= \sum_t \Pr(c \mid do(s), do(t))\Pr(t \mid s) && \text{(Rule 2)}\\
&= \sum_t \Pr(c \mid do(t))\Pr(t \mid s) && \text{(Rule 3)}\\
&= \sum_{s'} \sum_t \Pr(c \mid do(t), s')\Pr(s' \mid do(t))\Pr(t \mid s) && \text{(Probability Axioms)}
\end{aligned}
$$

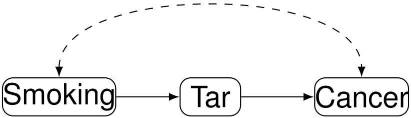
From do-calculus to front-door adjustment



What's the effect of smoking on cancer?

$$
\begin{aligned}
\Pr(c \mid do(s)) &= \sum_t \Pr(c \mid do(s), t) \Pr(t \mid do(s)) && \text{(Probability Axioms)} \\
&= \sum_t \Pr(c \mid do(s), do(t)) \Pr(t \mid do(s)) && \text{(Rule 2)} \\
&= \sum_t \Pr(c \mid do(s), do(t)) \Pr(t \mid s) && \text{(Rule 2)} \\
&= \sum_t \Pr(c \mid do(t)) \Pr(t \mid s) && \text{(Rule 3)} \\
&= \sum_{s'} \sum_t \Pr(c \mid do(t), s') \Pr(s' \mid do(t)) \Pr(t \mid s) && \text{(Probability Axioms)} \\
&= \sum_{s'} \sum_t \Pr(c \mid t, s') \Pr(s' \mid do(t)) \Pr(t \mid s) && \text{(Rule 2)}
\end{aligned}
$$

From do-calculus to front-door adjustment



What's the effect of smoking on cancer?

$$
\begin{aligned}
\Pr(c \mid do(s)) &= \sum_t \Pr(c \mid do(s), t) \Pr(t \mid do(s)) && \text{(Probability Axioms)} \\
&= \sum_t \Pr(c \mid do(s), do(t)) \Pr(t \mid do(s)) && \text{(Rule 2)} \\
&= \sum_t \Pr(c \mid do(s), do(t)) \Pr(t \mid s) && \text{(Rule 2)} \\
&= \sum_t \Pr(c \mid do(t)) \Pr(t \mid s) && \text{(Rule 3)} \\
&= \sum_{s'} \sum_t \Pr(c \mid do(t), s') \Pr(s' \mid do(t)) \Pr(t \mid s) && \text{(Probability Axioms)} \\
&= \sum_{s'} \sum_t \Pr(c \mid t, s') \Pr(s' \mid do(t)) \Pr(t \mid s) && \text{(Rule 2)} \\
&= \sum_{s'} \sum_t \Pr(c \mid t, s') \Pr(s') \Pr(t \mid s) && \text{(Rule 3)}
\end{aligned}
$$

From a calculus toward an automated algorithm

Limitations of the do-calculus:

- The do-calculus demands a lot of manual labor
- Non-identifiability is complicated

From a calculus toward an automated algorithm

Limitations of the do-calculus:

- The do-calculus demands a lot of manual labor
- Non-identifiability is complicated

Is it possible automatize it?

From a calculus toward an automated algorithm

Limitations of the do-calculus:

- The do-calculus demands a lot of manual labor
- Non-identifiability is complicated

Is it possible automatize it? Yes! There exists many algorithms.
In this course we will focus on the ID algorithm.

Table of content

Some lemmas

Lemma (adding do on non-ancestors)

If

$$\mathcal{W} = (\mathcal{V} \backslash \mathcal{X}) \backslash An(\mathcal{Y})_{\mathcal{G}_{\overline{X}}},$$

then

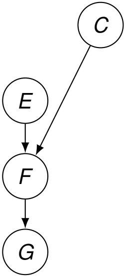
$$\Pr(y \mid do(x)) = \Pr(y \mid do(x), do(w)),$$

where w are arbatrary values of $\mathcal{W}$.

(proof on board)

Trees and forests

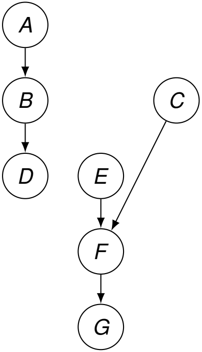
Tree A graph $\mathcal{G}$ such that each vertex has at most one child, and only one vertex (called the root) has no children.



Trees and forests

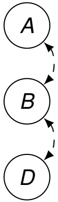
Tree A graph $\mathcal{G}$ such that each vertex has at most one child, and only one vertex (called the root) has no children.

Forest A graph $\mathcal{G}$ such that each vertex has at most one child.



C-components

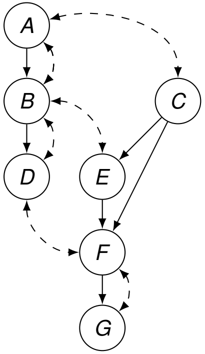
Confounded path A path where all directed arrowheads point at observable vertices, and never away from observable vertices.



C-components

Confounded path A path where all directed arrowheads point at observable vertices, and never away from observable vertices.

C-component A graph $\mathcal{G}$ where any pair of observable vertices is connected by a confounded path.



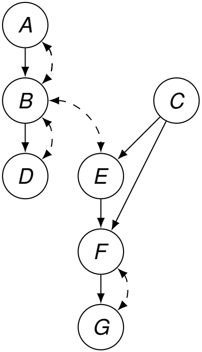
Decomposition into C-components

Any graph can be uniquely partitioned into a collection of subgraphs $C(\mathcal{G})$, each which is a maximal C-component.

Decomposition into C-components

Any graph can be uniquely partitioned into a collection of subgraphs $C(\mathcal{G})$, each which is a maximal C-component.

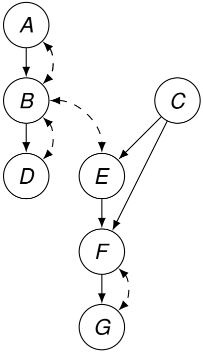
$$C(\mathcal{G}) = ?$$



Decomposition into C-components

Any graph can be uniquely partitioned into a collection of subgraphs $C(\mathcal{G})$, each which is a maximal C-component.

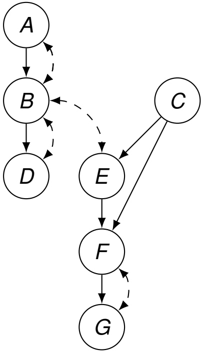
$$C(\mathcal{G}) = \begin{cases} \mathcal{G}[A, B, D, E] \\ \mathcal{G}[C] \\ \mathcal{G}[F, G] \end{cases}$$



Decomposition into C-components

Any graph can be uniquely partitioned into a collection of subgraphs $C(\mathcal{G})$, each which is a maximal C-component.

$$C(\mathcal{G}) = \begin{cases} \mathcal{G}[A, B, D, E] \\ \mathcal{G}[C] \\ \mathcal{G}[F, G] \end{cases}$$

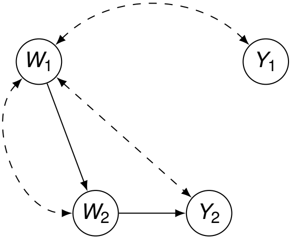


Lemma (c-component factorization) Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a causal graph. Let $C(\mathcal{G}\backslash\mathcal{X}) = \{\mathcal{S}_1, \cdots, \mathcal{S}_k\}$. Then

$$\Pr(y \mid do(x)) = \sum_{\mathcal{V}\backslash(y \cup x)} \prod_i \Pr(s_i \mid v\backslash s_i)$$

Proof in (Tian, 2002)

C-forest A graph $\mathcal{G}$ which is both a C-component and a forest. If a given C-forest has a set of root nodes $\mathcal{R}$, we call it $\mathcal{R}$-rooted.



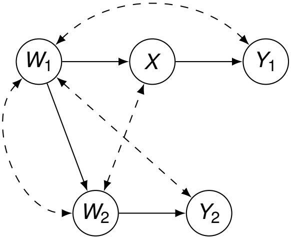
Hedges

C-forest A graph $\mathcal{G}$ which is both a C-component and a forest. If a given C-forest has a set of root nodes $\mathcal{R}$, we call it $\mathcal{R}$-rooted.

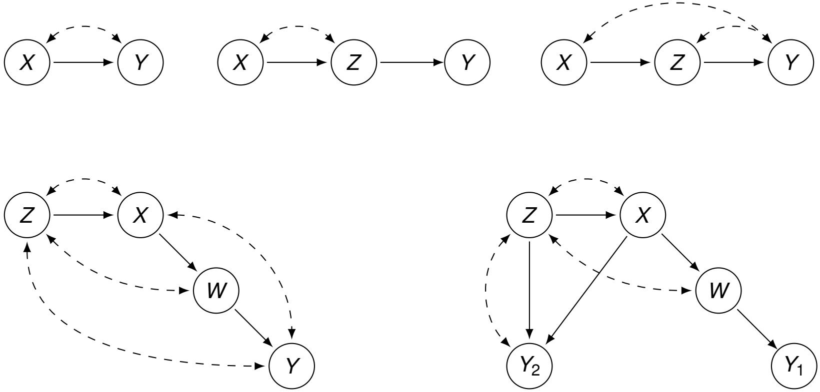
Hedge Let $\mathcal{X}, \mathcal{X} \in \mathcal{V}$ in $\mathcal{G}$. Let $\mathcal{H}, \mathcal{H}'$ be two R-rooted C-forests in $\mathcal{G}$ such that

- $\mathcal{H}' \subset \mathcal{H}$,
- $\mathcal{H} \cap X \neq \varnothing$,
- $\mathcal{H}' \cap X = \varnothing$, and
- $R \in An(Y)_{\mathcal{G}_{\underline{X}}}$.

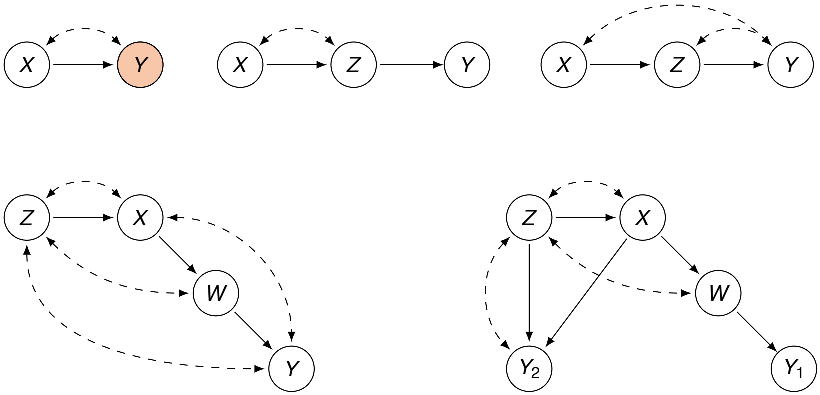
Then $\mathcal{H}$ and $\mathcal{H}'$ form a <u>hedge</u> for $P(y|do(x))$.



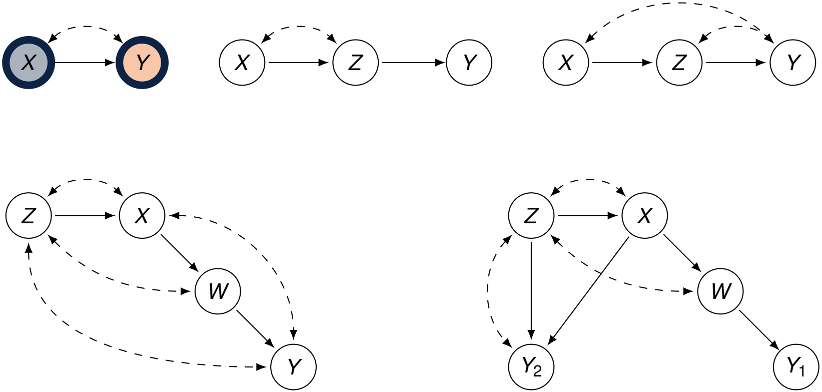
Find hedges for $\Pr(y \mid do(x))$



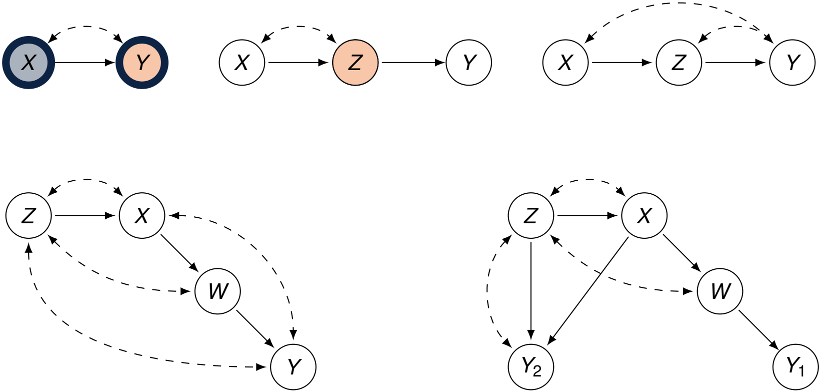
Find hedges for $\Pr(y \mid do(x))$



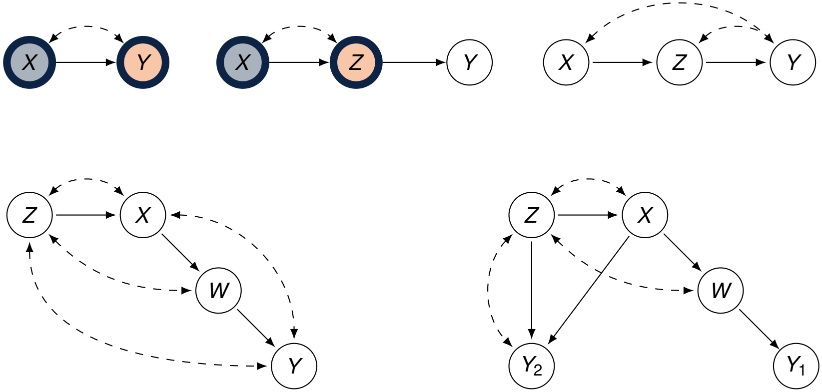
Find hedges for $\Pr(y \mid do(x))$



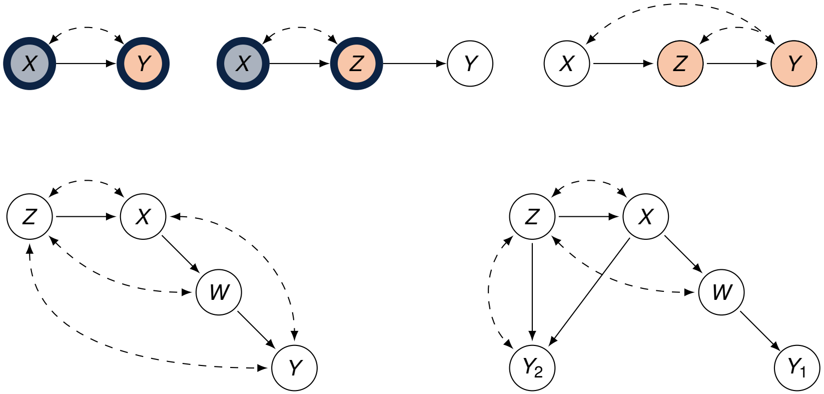
Find hedges for $\Pr(y \mid do(x))$



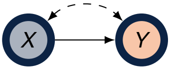
Find hedges for $\Pr(y \mid do(x))$

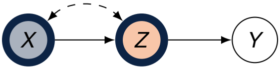


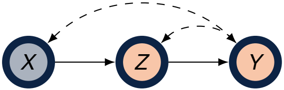
Find hedges for $\Pr(y \mid do(x))$

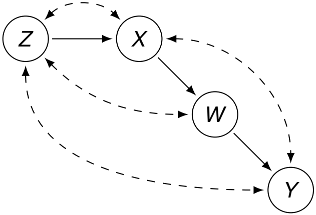


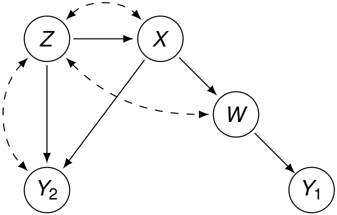
Find hedges for $\Pr(y \mid do(x))$



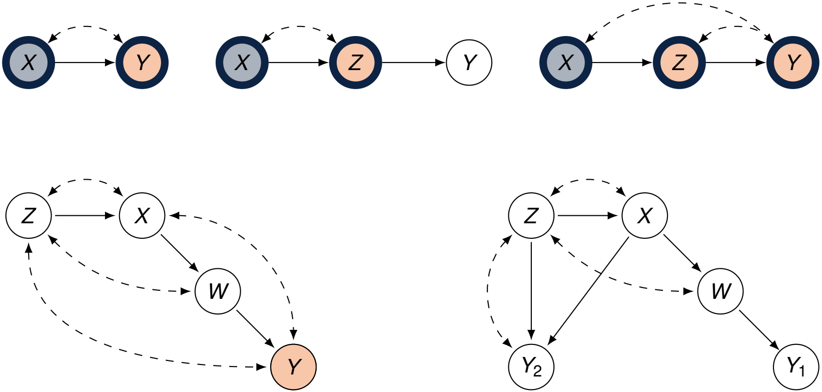




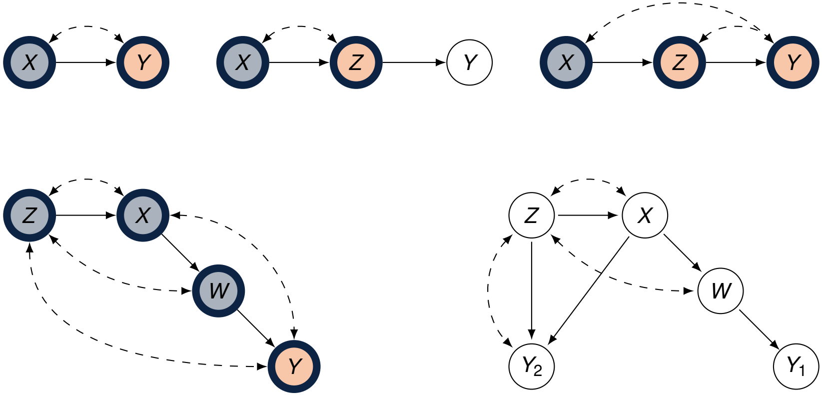




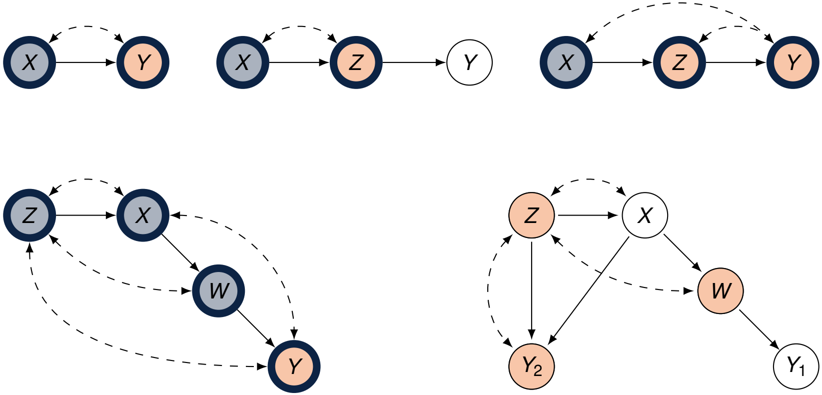
Find hedges for $\Pr(y \mid do(x))$



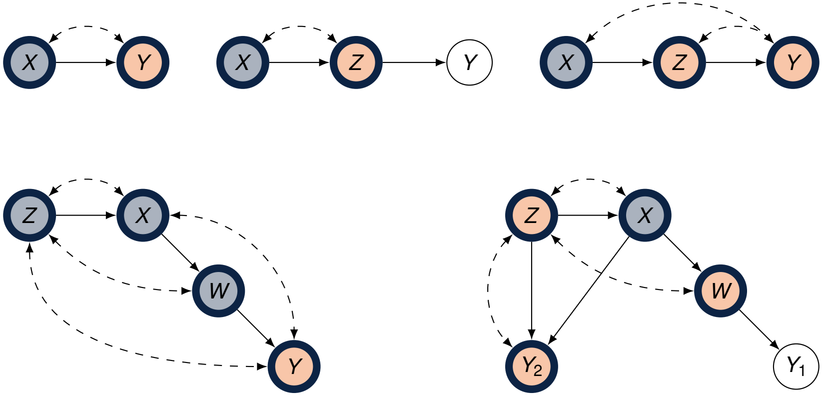
Find hedges for $\Pr(y \mid do(x))$



Find hedges for $\Pr(y \mid do(x))$



Find hedges for $\Pr(y \mid do(x))$



Hedges and non-identifiability

Theorem (Hedge criterion for non-identifiability) $\Pr(y \mid do(x))$ is not identifiable if and only if $\mathcal{G}$ contains a hedge for some $\Pr(y', do(x'))$, where $\mathcal{Y}' \in \mathcal{Y}$, $\mathcal{X}' \in \mathcal{X}$.

ID algorithm

Algorithm 1 ID

Input: $\mathcal{Y}, \mathcal{X}, \Pr(\mathcal{V}), \mathcal{G}$

Output: *do*-free expression for $\Pr(y \mid do(x))$ or FAIL$(\mathcal{H}, \mathcal{H}')$

if $\mathcal{X} = \varnothing$ **then**
    Return $\sum_{\mathcal{V}\backslash\mathcal{Y}} \Pr(v)$
if $\mathcal{V} \neq An(\mathcal{Y})_{\mathcal{G}}$ **then**
    Return ID$(y, x \cap An(\mathcal{Y})_{\mathcal{G}}, \sum_{\mathcal{V}\backslash An(\mathcal{Y})_{\mathcal{G}}} \Pr(v), \mathcal{G}[An(\mathcal{Y})_{\mathcal{G}}])$
if $\exists \mathcal{W} = (\mathcal{V}\backslash\mathcal{X})\backslash An(Y)_{\mathcal{G}_{\overline{X}}}$ such that $\mathcal{W} \neq \varnothing$ **then**
    Return ID$(y, x \cup w, P, \mathcal{G})$
if $C(\mathcal{G}\backslash\mathcal{X}) = \{\mathcal{S}_1, \cdots, \mathcal{S}_k\}$ (for $k \geq 2$) **then**
    Return $\sum_{\mathcal{V}\backslash(y \cup x)} \prod_i$ **ID**$(s_i, v\backslash s_i, \Pr(v), \mathcal{G})$
else if $C(\mathcal{G}\backslash\mathcal{X}) = \{\mathcal{S}\}$ **then**
    if $C(\mathcal{G}) = \{\mathcal{G}\}$ **then**
        Return FAIL$(\mathcal{G}, \mathcal{S})$
    if $S \in C(\mathcal{G})$ **then**
        Return $\sum_{s \setminus y} \prod_{V_i \in S} \Pr(v_i | v_{\pi}^{(i-1)})$
    if $\exists S', S \subseteq S' \in C(\mathcal{G})$ **then**
        Return ID$(y, x \cap S', \prod_{V_i \in S'} \Pr(V_i \mid V_{\pi}^{(i-1)} \cap S', v_{\pi}^{(i-1)}\backslash S'), S')$

ID algorithm

Algorithm 2 ID

Input: $\mathcal{Y}$, $\mathcal{X}$, $\Pr(\mathcal{V})$, $\mathcal{G}$

Output: *do*-free expression for $\Pr(y \mid do(x))$ or FAIL($\mathcal{H}, \mathcal{H}'$)

if $\mathcal{X} = \varnothing$ **then**
  Return $\sum_{\mathcal{V}\setminus\mathcal{Y}} \Pr(v)$
if $\mathcal{V} \neq An(\mathcal{Y})_{\mathcal{G}}$ **then**
  Return ID$(y, x \cap An(\mathcal{Y})_{\mathcal{G}}, \sum_{\mathcal{V}\setminus An(\mathcal{Y})_{\mathcal{G}}} \Pr(v), \mathcal{G}[An(\mathcal{Y})_{\mathcal{G}}])$
if $\exists \mathcal{W} = (\mathcal{V}\setminus\mathcal{X})\setminus An(Y)_{\mathcal{G}_{\overline{X}}}$ such that $\mathcal{W} \neq \varnothing$ **then**
  Return ID$(y, x \cup w, P, \mathcal{G})$
if $C(\mathcal{G}\setminus\mathcal{X}) = \{\mathcal{S}_1, \cdots, \mathcal{S}_k\}$ (for $k \geq 2$) **then**
  Return $\sum_{\mathcal{V}\setminus(y\cup x)} \prod_i$ **ID**$(s_i, v\setminus s_i, \Pr(v), \mathcal{G})$
else if $C(\mathcal{G}\setminus\mathcal{X}) = \{\mathcal{S}\}$ **then**
  if $C(\mathcal{G}) = \{\mathcal{G}\}$ **then**
    Return FAIL$(\mathcal{G}, \mathcal{S})$
  if $S \in C(\mathcal{G})$ **then**
    Return $\sum_{s\setminus y} \prod_{V_i \in S} \Pr(v_i | v_\pi^{(i-1)})$
  if $\exists S', S \subseteq S' \in C(\mathcal{G})$ **then**
    Return ID$(y, x \cap S', \prod_{V_i \in S'} \Pr(V_i \mid V_\pi^{(i-1)} \cap S', v_\pi^{(i-1)} \setminus S'), S')$

Trivial

Algorithm 3 ID

Input: $\mathcal{Y}$, $\mathcal{X}$, $\Pr(\mathcal{V})$, $\mathcal{G}$

Output: *do*-free expression for $\Pr(y \mid do(x))$ or FAIL$(\mathcal{H}, \mathcal{H}')$

```
if 𝒳 = ∅ then
    Return Σ_{𝒱\𝒴} Pr(v)
if 𝒱 ≠ An(𝒴)_𝒢 then
    Return ID(y, x ∩ An(𝒴)_𝒢, Σ_{𝒱\An(𝒴)_𝒢} Pr(v), 𝒢[An(𝒴)_𝒢])
if ∃𝒲 = (𝒱\𝒳)\An(Y)_{𝒢_X̄} such that 𝒲 ≠ ∅ then
    Return ID(y, x ∪ w, P, 𝒢)
if C(𝒢\𝒳) = {𝒮_1, ⋯, 𝒮_k} (for k ≥ 2) then
    Return Σ_{𝒱\(y∪x)} Π_i ID(s_i, v\s_i, Pr(v), 𝒢)
else if C(𝒢\𝒳) = {𝒮} then
    if C(𝒢) = {𝒢} then
        Return FAIL(𝒢, 𝒮)
    if S ∈ C(𝒢) then
        Return Σ_{s∖y} Π_{V_i∈S} Pr(v_i | v_π^{(i−1)})
    if ∃S′, S ⊆ S′ ∈ C(𝒢) then
        Return ID(y, x ∩ S′, Π_{V_i∈S′} Pr(V_i | V_π^{(i−1)} ∩ S′, v_π^{(i−1)}\S′), S′)
```

Trivial

ID algorithm

Algorithm 4 ID

Input: $\mathcal{Y}$, $\mathcal{X}$, $\Pr(\mathcal{V})$, $\mathcal{G}$

Output: *do*-free expression for $\Pr(y \mid do(x))$ or FAIL$(\mathcal{H}, \mathcal{H}')$

if $\mathcal{X} = \varnothing$ **then**
  Return $\sum_{\mathcal{V}\setminus\mathcal{Y}} \Pr(v)$
if $\mathcal{V} \neq An(\mathcal{Y})_{\mathcal{G}}$ **then**
  Return ID$(y, x \cap An(\mathcal{Y})_{\mathcal{G}}, \sum_{\mathcal{V}\setminus An(\mathcal{Y})_{\mathcal{G}}} \Pr(v), \mathcal{G}[An(\mathcal{Y})_{\mathcal{G}}])$
if $\exists \mathcal{W} = (\mathcal{V}\setminus\mathcal{X})\setminus An(Y)_{\mathcal{G}_{\overline{X}}}$ such that $\mathcal{W} \neq \varnothing$ **then**
  Return ID$(y, x \cup w, P, \mathcal{G})$
if $C(\mathcal{G}\setminus\mathcal{X}) = \{\mathcal{S}_1, \cdots, \mathcal{S}_k\}$ (for $k \geq 2$) **then**
  Return $\sum_{\mathcal{V}\setminus(y \cup x)} \prod_i$ **ID**$(s_i, v\setminus s_i, \Pr(v), \mathcal{G})$
else if $C(\mathcal{G}\setminus\mathcal{X}) = \{\mathcal{S}\}$ **then**
  if $C(\mathcal{G}) = \{\mathcal{G}\}$ **then**
    Return FAIL$(\mathcal{G}, \mathcal{S})$
  if $S \in C(\mathcal{G})$ **then**
    Return $\sum_{s \setminus y} \prod_{V_i \in S} \Pr(v_i | v_\pi^{(i-1)})$
  if $\exists S', S \subseteq S' \in C(\mathcal{G})$ **then**
    Return ID$(y, x \cap S', \prod_{V_i \in S'} \Pr(V_i \mid V_\pi^{(i-1)} \cap S', v_\pi^{(i-1)} \setminus S'), S')$

Lemma (adding do on non-ancestors)

ID algorithm

Algorithm 5 ID

Input: $\mathcal{Y}$, $\mathcal{X}$, $\Pr(\mathcal{V})$, $\mathcal{G}$

Output: *do*-free expression for $\Pr(y \mid do(x))$ or FAIL$(\mathcal{H}, \mathcal{H}')$

- **if** $\mathcal{X} = \emptyset$ **then**
  - **Return** $\sum_{\mathcal{V} \setminus \mathcal{Y}} \Pr(v)$
- **if** $\mathcal{V} \neq An(\mathcal{Y})_{\mathcal{G}}$ **then**
  - **Return ID**$(y, x \cap An(\mathcal{Y})_{\mathcal{G}}, \sum_{\mathcal{V} \setminus An(\mathcal{Y})_{\mathcal{G}}} \Pr(v), \mathcal{G}[An(\mathcal{Y})_{\mathcal{G}}])$
- **if** $\exists \mathcal{W} = (\mathcal{V} \setminus \mathcal{X}) \setminus An(Y)_{\mathcal{G}_{\overline{X}}}$ such that $\mathcal{W} \neq \emptyset$ **then**
  - **Return ID**$(y, x \cup w, P, \mathcal{G})$
- **if** $C(\mathcal{G} \setminus \mathcal{X}) = \{\mathcal{S}_1, \cdots, \mathcal{S}_k\}$ (for $k \geq 2$) **then**
  - **Return** $\sum_{\mathcal{V} \setminus (y \cup x)} \prod_i$ **ID**$(s_i, v \setminus s_i, \Pr(v), \mathcal{G})$
- **else if** $C(\mathcal{G} \setminus \mathcal{X}) = \{\mathcal{S}\}$ **then**
  - **if** $C(\mathcal{G}) = \{\mathcal{G}\}$ **then**
    - **Return FAIL**$(\mathcal{G}, \mathcal{S})$
  - **if** $S \in C(\mathcal{G})$ **then**
    - **Return** $\sum_{s \setminus y} \prod_{V_i \in S} \Pr(v_i | v_\pi^{(i-1)})$
  - **if** $\exists S', S \subseteq S' \in C(\mathcal{G})$ **then**
    - **Return ID**$(y, x \cap S', \prod_{V_i \in S'} \Pr(V_i \mid V_\pi^{(i-1)} \cap S', v_\pi^{(i-1)} \setminus S'), S')$

Lemma (c-component factorization)

ID algorithm

Algorithm 6 ID

Input: $\mathcal{Y}$, $\mathcal{X}$, $\Pr(\mathcal{V})$, $\mathcal{G}$

Output: *do*-free expression for $\Pr(y \mid do(x))$ or FAIL$(\mathcal{H}, \mathcal{H}')$

if $\mathcal{X} = \varnothing$ **then**

  Return $\sum_{\mathcal{V}\setminus\mathcal{Y}} \Pr(v)$

if $\mathcal{V} \neq An(\mathcal{Y})_{\mathcal{G}}$ **then**

  Return ID$(y, x \cap An(\mathcal{Y})_{\mathcal{G}}, \sum_{\mathcal{V}\setminus An(\mathcal{Y})_{\mathcal{G}}} \Pr(v), \mathcal{G}[An(\mathcal{Y})_{\mathcal{G}}])$

if $\exists \mathcal{W} = (\mathcal{V}\setminus\mathcal{X})\setminus An(Y)_{\mathcal{G}_{\overline{X}}}$ such that $\mathcal{W} \neq \varnothing$ **then**

  Return ID$(y, x \cup w, P, \mathcal{G})$

if $C(\mathcal{G}\setminus\mathcal{X}) = \{\mathcal{S}_1, \cdots, \mathcal{S}_k\}$ (for $k \geq 2$) **then**

  Return $\sum_{\mathcal{V}\setminus(y\cup x)} \prod_i$ **ID**$(s_i, v\setminus s_i, \Pr(v), \mathcal{G})$

else if $C(\mathcal{G}\setminus\mathcal{X}) = \{\mathcal{S}\}$ **then**

  if $C(\mathcal{G}) = \{\mathcal{G}\}$ **then**

    Return FAIL$(\mathcal{G}, \mathcal{S})$

  if $S \in C(\mathcal{G})$ **then**

    Return $\sum_{s\setminus y} \prod_{V_i \in S} \Pr(v_i | v_{\pi}^{(i-1)})$

  if $\exists S', S \subseteq S' \in C(\mathcal{G})$ **then**

    Return ID$(y, x \cap S', \prod_{V_i \in S'} \Pr(V_i \mid V_{\pi}^{(i-1)} \cap S', v_{\pi}^{(i-1)} \setminus S'), S')$

ID algorithm

Algorithm 7 ID

Input: $\mathcal{Y}$, $\mathcal{X}$, $\Pr(\mathcal{V})$, $\mathcal{G}$

Output: *do*-free expression for $\Pr(y \mid do(x))$ or $\text{FAIL}(\mathcal{H}, \mathcal{H}')$

1: **if** $\mathcal{X} = \varnothing$ **then**
2: **Return** $\sum_{\mathcal{V}\setminus\mathcal{Y}} \Pr(v)$
3: **if** $\mathcal{V} \neq An(\mathcal{Y})_{\mathcal{G}}$ **then**
4: **Return** **ID**$(y, x \cap An(\mathcal{Y})_{\mathcal{G}}, \sum_{\mathcal{V}\setminus An(\mathcal{Y})_{\mathcal{G}}} \Pr(v), \mathcal{G}[An(\mathcal{Y})_{\mathcal{G}}])$
5: **if** $\exists \mathcal{W} = (\mathcal{V}\setminus\mathcal{X})\setminus An(Y)_{\mathcal{G}_{\overline{X}}}$ such that $\mathcal{W} \neq \varnothing$ **then**
6: **Return** **ID**$(y, x \cup w, P, \mathcal{G})$
7: **if** $C(\mathcal{G}\setminus\mathcal{X}) = \{\mathcal{S}_1, \cdots, \mathcal{S}_k\}$ (for $k \geq 2$) **then**
8: **Return** $\sum_{\mathcal{V}\setminus(y \cup x)} \prod_i$ **ID**$(s_i, v\setminus s_i, \Pr(v), \mathcal{G})$
9: **else if** $C(\mathcal{G}\setminus\mathcal{X}) = \{\mathcal{S}\}$ **then**
10: **if** $C(\mathcal{G}) = \{\mathcal{G}\}$ **then**
11: **Return FAIL**$(\mathcal{G}, \mathcal{S})$
12: **if** $S \in C(\mathcal{G})$ **then**
13: **Return** $\sum_{s \setminus y} \prod_{V_i \in S} \Pr(v_i | v_{\pi}^{(i-1)})$
14: **if** $\exists S', S \subseteq S' \in C(\mathcal{G})$ **then**
15: **Return** **ID**$(y, x \cap S', \prod_{V_i \in S'} \Pr(V_i \mid V_{\pi}^{(i-1)} \cap S', v_{\pi}^{(i-1)} \setminus S'), S')$

Theorem (Hedge criterion for non-identifiability)

ID algorithm

Algorithm 8 ID

Input: $\mathcal{Y}$, $\mathcal{X}$, $\Pr(\mathcal{V})$, $\mathcal{G}$

Output: *do*-free expression for $\Pr(y \mid do(x))$ or FAIL$(\mathcal{H}, \mathcal{H}')$

if $\mathcal{X} = \emptyset$ **then**
  Return $\sum_{\mathcal{V} \setminus \mathcal{Y}} \Pr(v)$
if $\mathcal{V} \neq An(\mathcal{Y})_{\mathcal{G}}$ **then**
  Return ID$(y, x \cap An(\mathcal{Y})_{\mathcal{G}}, \sum_{\mathcal{V} \setminus An(\mathcal{Y})_{\mathcal{G}}} \Pr(v), \mathcal{G}[An(\mathcal{Y})_{\mathcal{G}}])$
if $\exists \mathcal{W} = (\mathcal{V} \setminus \mathcal{X}) \setminus An(Y)_{\mathcal{G}_{\overline{X}}}$ such that $\mathcal{W} \neq \emptyset$ **then**
  Return ID$(y, x \cup w, P, \mathcal{G})$
if $C(\mathcal{G} \setminus \mathcal{X}) = \{\mathcal{S}_1, \cdots, \mathcal{S}_k\}$ (for $k \geq 2$) **then**
  Return $\sum_{\mathcal{V} \setminus (y \cup x)} \prod_i$ **ID**$(s_i, v \setminus s_i, \Pr(v), \mathcal{G})$
else if $C(\mathcal{G} \setminus \mathcal{X}) = \{\mathcal{S}\}$ **then**
  if $C(\mathcal{G}) = \{\mathcal{G}\}$ **then**
    Return FAIL$(\mathcal{G}, \mathcal{S})$
  if $S \in C(\mathcal{G})$ **then**
    Return $\sum_{s \setminus y} \prod_{V_i \in S} \Pr(v_i | v_{\pi}^{(i-1)})$
  if $\exists S', S \subseteq S' \in C(\mathcal{G})$ **then**
    Return ID$(y, x \cap S', \prod_{V_i \in S'} \Pr(V_i \mid V_{\pi}^{(i-1)} \cap S', v_{\pi}^{(i-1)} \setminus S'), S')$

Proof in (Shpitser and Pearl, 2006)

ID algorithm

Algorithm 9 ID

Input: $\mathcal{Y}$, $\mathcal{X}$, $\Pr(\mathcal{V})$, $\mathcal{G}$

Output: *do*-free expression for $\Pr(y \mid do(x))$ or FAIL$(\mathcal{H}, \mathcal{H}')$

if $\mathcal{X} = \emptyset$ **then**
    Return $\sum_{\mathcal{V}\setminus\mathcal{Y}} \Pr(v)$
if $\mathcal{V} \neq An(\mathcal{Y})_{\mathcal{G}}$ **then**
    Return ID$(y, x \cap An(\mathcal{Y})_{\mathcal{G}}, \sum_{\mathcal{V}\setminus An(\mathcal{Y})_{\mathcal{G}}} \Pr(v), \mathcal{G}[An(\mathcal{Y})_{\mathcal{G}}])$
if $\exists \mathcal{W} = (\mathcal{V}\setminus\mathcal{X})\setminus An(Y)_{\mathcal{G}_{\overline{X}}}$ such that $\mathcal{W} \neq \emptyset$ **then**
    Return ID$(y, x \cup w, P, \mathcal{G})$
if $C(\mathcal{G}\setminus\mathcal{X}) = \{\mathcal{S}_1, \cdots, \mathcal{S}_k\}$ (for $k \geq 2$) **then**
    Return $\sum_{\mathcal{V}\setminus(y\cup x)} \prod_i$ **ID**$(s_i, v\setminus s_i, \Pr(v), \mathcal{G})$
else if $C(\mathcal{G}\setminus\mathcal{X}) = \{\mathcal{S}\}$ **then**
    if $C(\mathcal{G}) = \{\mathcal{G}\}$ **then**
        Return FAIL$(\mathcal{G}, \mathcal{S})$
    if $S \in C(\mathcal{G})$ **then**
        Return $\sum_{s\setminus y} \prod_{V_i \in S} \Pr(v_i | v_\pi^{(i-1)})$
    if $\exists S', S \subseteq S' \in C(\mathcal{G})$ **then**
        Return ID$(y, x \cap S', \prod_{V_i \in S'} \Pr(V_i \mid V_\pi^{(i-1)} \cap S', v_\pi^{(i-1)}\setminus S'), S')$

Proof in (Shpitser and Pearl, 2006)

Completeness of ID algorithm

Theorem (Soudness of the ID algorithm) Whenever the ID algorithm returns an expression for $\Pr(y \mid do(x))$, it is correct.

Partially proved in the previous slides.

Completeness of ID algorithm

Theorem (Soudness of the ID algorithm) Whenever the ID algorithm returns an expression for $\Pr(y \mid do(x))$, it is correct.

Partially proved in the previous slides.

Theorem (Completeness of ID algorithm) ID is complete.

Proof in (Shpitser and Pearl, 2006)

Table of content

Conclusion

- do calculus is complete;
- The ID algorithm is complete.

Conclusion

- do calculus is complete;
- The ID algorithm is complete.

Some extensions

- The IDC algorithm that support conditioning;
- Finding optimal adjustment sets;
- Identifiability for direct effects and indirect effects.

Some extensions

- The IDC algorithm that support conditioning;
- Finding optimal adjustment sets;
- Identifiability for direct effects and indirect effects.

Some extensions

- The IDC algorithm that support conditioning;
- Finding optimal adjustment sets;
- Identifiability for direct effects and indirect effects.

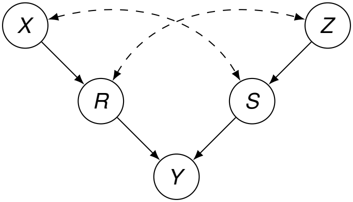
References

Direct inspirations

1. *Causal diagrams for empirical research*, J. Pearl. Biometrika, 1995
2. *Identification of Joint Interventional Distributions in Recursive Semi-Markovian Causal Models*, I. Shpitser, J. Pearl. Proceedings of the Twenty National Conference on Artificial Intelligence, 2006
3. *Complete Identification Methods for the Causal Hierarchy*, I. Shpitser, J. Pearl. Journal of Machine Learning Research, 2008
4. *Studies in Causal Reasoning and Learning*, J. Tian. PhD thesis, 2002
5. *Causality*, J. Pearl. Cambridge University Press, 2nd edition, 2009

Exercise 1.1

Consider the following semi-Markovian model:



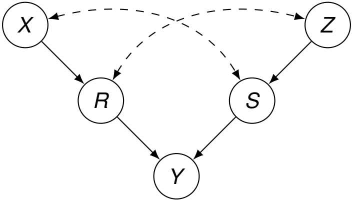
Test, using do-calculus, whether the causal effect

$$P(y \mid do(r))$$

is identifiable. If the answer is yes, provide an expression for it that does not contain the *do*() operator.

Exercise 1.2

Consider the following semi-Markovian model:



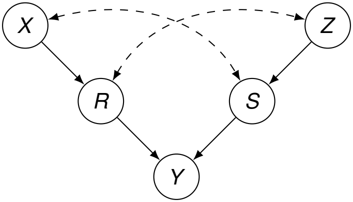
Test, using do-calculus, whether the causal effect

$$P(r \mid do(y))$$

is identifiable. If the answer is yes, provide an expression for it that does not contain the *do*() operator.

Exercise 1.3

Consider the following semi-Markovian model:



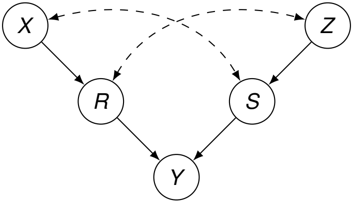
Test, using do-calculus, whether the causal effect

$$P(y \mid do(r), do(s))$$

is identifiable. If the answer is yes, provide an expression for it that does not contain the *do*() operator.

Exercise 1.4

Consider the following semi-Markovian model:



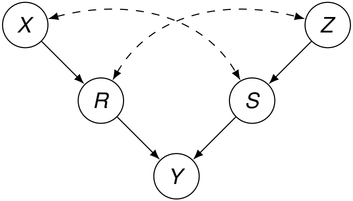
Test, using do-calculus, whether the causal effect

$$P(r \mid do(x), do(z))$$

is identifiable. If the answer is yes, provide an expression for it that does not contain the *do*() operator.

Exercise 1.5

Consider the following semi-Markovian model:



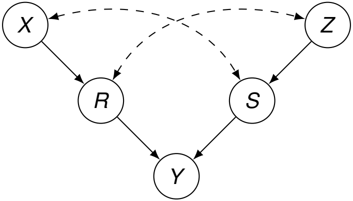
Test, using do-calculus, whether the causal effect

$$P(s \mid do(x), do(z))$$

is identifiable. If the answer is yes, provide an expression for it that does not contain the *do*() operator.

Exercise 1.6

Consider the following semi-Markovian model:



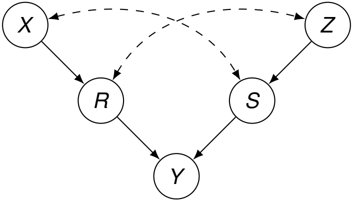
Test, using do-calculus, whether the causal effect

$$P(r, s \mid do(x), do(z))$$

is identifiable. If the answer is yes, provide an expression for it that does not contain the *do*() operator.

Exercise 1.7

Consider the following semi-Markovian model:



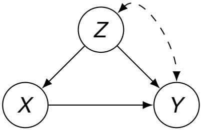
Test, using do-calculus, whether the causal effect

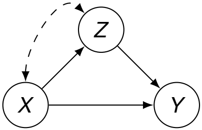
$$P(y \mid do(x), do(z))$$

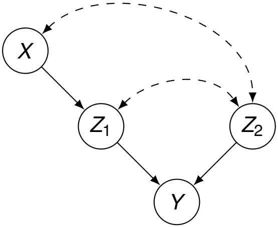
is identifiable. If the answer is yes, provide an expression for it that does not contain the *do*() operator.

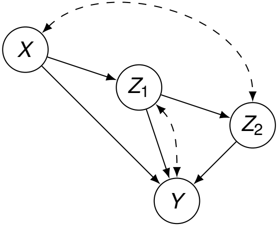
Exercise 2

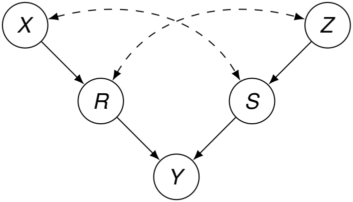
Which of the following semi-Markovian models admet an identifiable causal effect $\Pr(y \mid do(x))$?





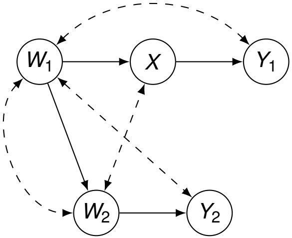






Exercise 3

Consider the following semi-Markovian model containing a hedge for $\Pr(y \mid do(x))$:



- Is it possible to remove the hedge by adding one directed edge to the graph? If yes, which one?
- Is it possible to remove the hedge by deleting one directed edge from the graph? If yes, which one?